ATTACHMENT

All correspondence and telephonic communications relating to this patent application should be directed to:

RICK D. NYDEGGER
WORKMAN, NYDEGGER & SEELEY
1000 Eagle Gate Tower
60 East South Temple
Salt Lake City, Utah 84111
Telephone: (801) 533-9800
Facsimile: (801) 328-1707

Express Mailing Label No. EL 813 876 742 US

PATENT APPLICATION
Docket No. 14531.134

UNITED STATES PATENT APPLICATION

of

Shankar Moni

and

John Tardif

for

SPATIALLY TRANSCODING A VIDEO STREAM

WORKMAN, NYDEGGER & SEELEY
A PROFESSIONAL CORPORATION
ATTORNEYS AT LAW
1000 EAGLE GATE TOWER
60 EAST SOUTH TEMPLE
SALT LAKE CITY, UTAH 84111

SPATIALLY TRANSCODING A VIDEO STREAM

BACKGROUND OF THE INVENTION

The Field of the Invention

[01] The present invention relates to reducing an image size and/or a bit rate of a video stream. More particularly, the present invention relates to reducing a bit rate of a video stream by spatially transcoding the video stream.

Background and Relevant Art

[02] Digital video signals have several significant advantages over their analog counterparts. They can be transmitted, for example, over long distances and stored without degradation. One cost, however, of digital signals is related to the bandwidth that they consume. The raw storage requirement for a typical uncompressed video stream, depending on the resolution, is approximately 20 megabytes per second. At this rate, an uncompressed two hour movie would require 144 Gigabytes of memory, well above the capacity of a conventional Digital Versatile Disk (DVD).

[03] Clearly, uncompressed digital or video signals can consume significant bandwidth. This is particularly a problem in situations where multiple digital signals are being broadcast or where the digital signal is being transmitted over a medium such as the Internet, which has limited bandwidth in many circumstances. The need to reduce bandwidth requirements of digital signals such as video streams has led to the development of various compression schemes.

[04] One conventional compression scheme or standard defined by the Moving Pictures Expert Group (MPEG) is called MPEG-2. MPEG-2 is based on the principle that there is a

WORKMAN, NYDEGGER & SEELEY
A PROFESSIONAL CORPORATION
ATTORNEYS AT LAW
1000 EAGLE GATE TOWER
60 EAST SOUTH TEMPLE
SALT LAKE CITY, UTAH 84111

large degree of visual redundancy in video streams and that video storage and bandwidth requirements can be reduced by removing the redundant information from the video stream.

[05] The bit stream defined by MPEG is the output of an encoding process that is designed to significantly compress the video picture information. As the MPEG standard only defines the syntax of the resulting bit stream, the standard is flexible enough to be used in a variety of different situations, such as satellite broadcast services, cable television, interactive television services, and the Internet.

[06] The MPEG encoding process generally occurs as follows. A video signal is sampled and quantized to define color and luminance components for each pixel of the digital video. Values representing the color and luminance components are stored in structures known as macroblocks. The color and luminance values stored in the macroblocks are converted to frequency values using a discrete cosine transform (DCT). The transform coefficients obtained from the DCT represent different frequencies in the brightness and the color of the picture.

[07] The MPEG encoding process takes advantage of the fact that human visual system is insensitive to high frequencies in color and luminance changes, and quantizes the transform coefficients to represent the color and luminance information by smaller or more coarsely scaled values. The quantized DCT transform coefficients are then encoded using run level coding (RLC) and variable length coding (VLC) techniques, which further compress the video stream.

[08] The MPEG standard also provides additional compression through motion compensation techniques. Under the MPEG standards, there are three types of pictures or frames: I frames; P frames; and B frames. The I frames are intra-coded, meaning that they can be reconstructed without reference to any other frame or picture in the video stream. P

frames and B frames are inter-coded, meaning that they are reconstructed by referencing another frame or picture. For example, P frames and B frames contain motion vectors that represent estimated motion with respect to the reference frame(s). The use of motion vectors enables an MPEG encoder to reduce the bandwidth requirements of a particular video stream.

[09] However, even compressed video or MPEG streams may still have a bit rate that is unsatisfactorily high for certain applications, and therefore there is a need to further compress the video stream or reduce the bit rate of the stream. One solution to this problem is to reduce the bit rate of the MPEG stream by completely decoding the MPEG stream and then re-encoding the MPEG stream at a higher compression ratio to reduce the bit rate to an acceptable level. However, decoding and re-encoding an MPEG stream in this fashion is often computationally expensive because of the need to perform an inverse quantization and an inverse DCT to recreate an approximation of the original data prior to re-encoding the data in accordance with a desired bit rate. There is also a need to recompute motion vectors and other parameters that are included in the resulting bit stream.

[010] As previously stated however, decoding and re-encoding a video stream is often necessary because the bit rate of the incoming video stream may be higher than the available bandwidth or the bit rate of the incoming video stream may be higher than the optimal bit rate for storage of the video stream on a storage medium such as a hard disk drive. In view of these and other problems presented by video streams, minimally complex systems and methods are needed that can reduce the storage and bandwidth requirements of a video stream.

SUMMARY OF THE INVENTION

[011] The present invention recognizes the limitations of the prior art and the need for systems, methods, and computer program products that are able to reduce the bit rate of a video stream. Reducing the bit rate of a video stream provides significant advantages such as reducing bandwidth and storage requirements of a video stream, enabling the viewing of high definition video streams on a standard definition device, and allowing users to store or render video streams at bit rates and image sizes that are determined by the user.

[012] Reducing the bit rate of a video stream begins by decoding the video stream. After the video stream has been decoded, each image of the video stream is resized or spatially reduced horizontally and vertically by a factor. The horizontal and vertical scaling factors may be different. After the images have been resized, the outgoing video stream is generated. Instead of re-encoding the video stream from the decoded video stream, the present invention utilizes parameters that were part of or that described the original incoming video stream. These parameters represent decisions made by a previous encoder that more accurately reflect the video stream. The video stream generator thus utilizes these parameters as the new video stream is generated instead of generating the parameters strictly from the decoded video stream.

[013] In some instances, some of these parameters from the original video stream are unchanged in the transcoded video stream while other parameters of the transcoded video stream are re-computed. Re-computing a particular parameter is often necessary, for instance, because the spatial size of the images has changed. Motion vectors, in particular, are re-computed to account for the changed image size. Other macroblock parameters are also re-computed using a variety of procedures that take the parameter values of the original video stream into account. This results in a video stream that is representative of the

WORKMAN, NYDEGGER & SEELEY
A PROFESSIONAL CORPORATION
ATTORNEYS AT LAW
1000 EAGLE GATE TOWER
60 EAST SOUTH TEMPLE
SALT LAKE CITY, UTAH 84111

sequenced reduced images, has a reduced bit rate, retains improved visual quality, and whose generation is computationally efficient.

[014] Additional features and advantages of the invention will be set forth in the description which follows, and in part will be obvious from the description, or may be learned by the practice of the invention. The features and advantages of the invention may be realized and obtained by means of the instruments and combinations particularly pointed out in the appended claims. These and other features of the present invention will become more fully apparent from the following description and appended claims, or may be learned by the practice of the invention as set forth hereinafter.

BRIEF DESCRIPTION OF THE DRAWINGS

[015] In order to describe the manner in which the above-recited and other advantages and features of the invention can be obtained, a more particular description of the invention briefly described above will be rendered by reference to specific embodiments thereof which are illustrated in the appended drawings. Understanding that these drawings depict only typical embodiments of the invention and are not therefore to be considered to be limiting of its scope, the invention will be described and explained with additional specificity and detail through the use of the accompanying drawings in which:

[016] Figure 1 illustrates exemplary levels of a typical video stream and the headers of those levels, wherein the headers contain stream parameters.

[017] Figure 2 is a block diagram illustrating a spatial transcoder that receives an incoming video stream and transcodes the incoming video stream to generate a spatially reduced video stream;

[018] Figure 3 is a more detailed block diagram of a spatial transcoder and illustrates how stream parameters are used to generate a new video stream;

[019] Figure 4 is a block diagram that illustrates an example of how macroblocks of the incoming video stream are mapped to the outgoing video stream such that parameters of the new macroblocks can be generated from the parameters of the original macroblocks;;

[020] Figure 5A is a block diagram illustrating a raster scan ordering of pixels;

[021] Figure 5B illustrates the order in which the pixels of Figure 5A were selected for subsampled sum of absolute differences;

[022] Figure 6 is a block diagram illustrating how Discrete Cosine Transfer coefficients are generated by a stream generator; and

[023] Figure 7 is a block diagram illustrating a suitable operating environment for the present invention.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

[024] An uncompressed digital video stream has high bandwidth and storage requirements. Video streams are encoded in order to reduce the number of bits that must be transmitted and thus reduce the bandwidth and storage requirements of the video stream. In some situations, however, it is necessary to further reduce the number of bits in a video stream for various reasons, including bandwidth restrictions and storage concerns.

[025] The present invention relates to transcoding a video stream such as a Moving Pictures Experts Group (MPEG) stream. More specifically, the present invention relates to spatial transcoding where the outgoing video stream that has been generated from the incoming video stream has images that are spatially smaller than the images in the original video stream.

[026] Video streams received over satellite systems, cable systems, and the Internet, for example, have already been encoded. In some situations, a high quality encoder was utilized to generate the video stream. The parameters of the video stream (motion type, width, height, picture rate, bit rate, etc.) represent decisions that are made by the encoder and are used as the video stream is transcoded.

[027] As previously mentioned, transcoding a video stream is often performed by fully decoding the original video stream and then re-encoding the video stream. Transcoding a video stream in this fashion ignores the parameters of the original video stream. In other words, the parameters of the new video stream will be regenerated or recomputed without reference to the original parameters. In addition to being computationally expensive, the encoding decisions of the original encoder are therefore not considered as the video stream is re-encoded. In contrast, the present invention generates a new video stream using the parameters or decisions that were made during the previous encoding of the video stream.

WORKMAN, NYDEGGER & SEELEY
A PROFESSIONAL CORPORATION
ATTORNEYS AT LAW
1000 EAGLE GATE TOWER
60 EAST SOUTH TEMPLE
SALT LAKE CITY, UTAH 84111

This ensures that the transcoded video stream more closely approximates the original video stream and yields a substantially smaller implementation size.

[028] An exemplary use of the present invention is the ability to view a high definition video stream on a standard definition device. A high definition video stream, in addition to having a relatively high bit rate, contains images that cannot be shown on a standard definition device for various reasons, including image size. The present invention is able to spatially reduce the size of the images and reduce the bit rate of the video stream, thus enabling a standard definition device to display a video stream that was originally a high definition video stream.

[029] Figure 1 describes an exemplary video stream that has been encoded. Figure 1 also introduces some of the parameters that are included in an encoded video stream. These parameters often represent decisions made by the previous encoder during the encoding process. Figure 1 is not intended as an exhaustive explanation of a video stream and it is further understood that the principles described herein can be applied more broadly to other parameters of the video stream.

[030] In this example, the video stream of Figure 1 illustrates a nested hierarchy of different levels of a video stream 99 (not all levels of a video stream are illustrated) that includes sequences, groups of pictures, pictures, slices, and macroblocks. Each subsequent level in the video stream is part of a previous layer or level. Thus the sequence level 101 is a series of sequences and each sequence contains or more groups of pictures (GOP). The group of picture (GOP) level is a series of groups of pictures and each GOP includes one or more pictures. The picture level 109 is a series of pictures (including I frames, P frames, and/or B frames) and each picture includes one or more slices. The slice level 113 is a series

of slices and each slice contains one or more macroblocks. The macroblock (MB) level 119 is a series of macroblocks.

[031] In order to decode the video stream 99, it is necessary to have certain information about the video stream 99. Often, this information is included in headers that are included in the video stream. Thus, each block of data at each level of the video stream usually has a header that contains relevant information that is related to the encoding and decoding of the video stream. The sequence 100, for example has a sequence header 102. The GOP 104 has a GOP header 106, the picture 108 has a picture header 110, the slice 112 has a slice header 114, and the MB 118 has a MB header 120.

[032] The sequence header 102 includes parameters that describe, for example, the width of pictures, the height of pictures, the aspect ratio of pixels, the picture rate, and the like. The sequence header 102 also includes parameters for the bit rate, the buffer size, and other flags. The sequence header 102 is also used to transmit the quantizer matrices for intra blocks and non intra blocks, for example.

[033] The GOP header 106 includes parameters that relate to a time code and other parameters that describe the structure of the GOP. The picture header 110 includes parameters that describe the type of frame or picture (I frame, P frame, and B frame in MPEG), for example. The picture parameters also include buffer parameters that indicate when decoding should begin and encode parameters that indicate whether half pixel motion vectors were used. The slice header 114 includes parameters to indicate which line the slice starts on and a quantizer scale indicating how the quantization table should be scaled for a particular slice. The MB header 120 indicates whether the MB 118 includes motion vectors as well as the type of motion vector (forward, backwards), type of macroblock, a

quantization scale, and the like. The MB header 120 also determines discrete cosine transform (DCT) type, DCT coefficients, coded block pattern, and associated flags.

[034] Figure 2 is a block diagram illustrating the functionality of the present invention. An incoming video stream 200 is received by a spatial transcoder 204. The spatial transcoder 204 decodes the incoming video stream 200 and resamples the video stream 200. From the resampled data and using parameters or cues from the incoming video stream 200, an outgoing video stream 202 is generated.

[035] The outgoing video stream 202 is spatially smaller because the images have been reduced in size. The outgoing video stream 202 also has a reduced bit rate compared to the incoming video stream 200. The spatial transcoder 204 does not, however, perform a full decode and a complete re-encoding of the video stream. The spatial transcoder 204 utilizes the decisions of the encoder that encoded the incoming video stream 200 in generating the outgoing video stream 202. The spatial transcoder 204 thus generates a video stream that corresponds to a reduced size image sequence that often results in a substantial bit rate reduction.

[036] Figure 3 is a block diagram that more fully illustrates the spatial transcoder of Figure 2. The input video stream 301 is received by a stream decoder 302, which decodes the video stream. Typically, the stream decoder 302 fully decodes the input video stream 301. The stream parameters 308 from the input video stream are extracted and saved for later use by the spatial transcoder 300 in generating the transcoded video stream. The stream parameters 308, as previously described, correspond to decisions of the encoder that encoded the input video stream 301. The spatial transcoder 300, by using these stream parameters in generating the output stream 307, is able to preserve those decisions, which often helps retain visual quality of the output video stream 307.

WORKMAN, NYDEGGER & SEELEY
A PROFESSIONAL CORPORATION
ATTORNEYS AT LAW
1000 EAGLE GATE TOWER
60 EAST SOUTH TEMPLE
SALT LAKE CITY, UTAH 84111

[037] The video stream is then resampled by the resampler 304 in order to reduce the image size. After the images in the video stream have been spatially reduced, the stream generator 306 generates the output video stream 307. It is not necessary for all images to be spatially reduced before the stream generator 306 begins generating the output video stream 307. As previously stated, the resampler 304 and the stream generator 306 utilize the stream parameters 308 from the input stream 301. To fully generate the output video stream 307, especially B frames and P frames, reference images 310 are made available to the stream generator 306 by a stream decoder 309, which decodes images from the output video stream 307.

[038] At the sequence level, GOP level, and picture level of the video stream, some of the parameters of the input video stream 301 may be altered. However, any change made to the parameters usually uses the original parameters as a reference for altering the parameters. In other words, the parameters of the output video stream 307 are related to the stream parameters 308 and are not strictly derived from the decoded video stream.

[039] At the sequence level, for example, a new picture size is computed as the output stream 307 has been spatially reduced by the spatial transcoder 300. Although the horizontal and vertical re-sampling factors used by the resampler 304 are already known by the spatial transcoder 300, it is useful to ensure that the height and width of the images in the output video stream are multiples of 32 and 16 respectively.

[040] At the picture level of the video stream, there is an f_code parameter included in the picture level header. The f_codes determine the granularity at which the motion vectors are encoded. This has a direct effect on the number of bits that are used to encode the residue of the motion vectors. The maximum motion vector of the picture determines the optimal or smallest f_code that can be used. In one example, the f_codes are scaled as they are

decoded. In another example, the f_codes are determined at the end of decoding the original picture.

[041] At the macroblock level of the video stream, there are several macroblock parameters than need to be determined. These include, but are not limited to, macroblock type, motion vectors and associated flags, DCT type, quantizer scale, coded block pattern, and DCT coefficients. The following paragraphs illustrate how these parameters may be determined by the spatial transcoder 300 or by the stream generator 306.

I. Macroblock Type

[042] Figure 4 illustrates a pair of pictures or frames. The picture 402 is present in the original video stream and the picture 404 is the transcoded version of the picture 402. As illustrated, the picture 404 has been spatially reduced in comparison to picture 402. The pictures 402, and 404 can correspond to various types of pictures, for example an I frame, a B frame or a P frame.

[043] The picture 402 of Figure 4 illustrates various macroblocks numbered as macroblock 405 through macroblock 412. As the picture 402 is transcoded, one or more of the macroblocks of the picture 402 are mapped or correspond to a macroblock of the picture 404. If the horizontal and/or vertical factors by which the picture 402 is scaled are not an integer value, it is possible that a partial macroblock in the picture 402 will be mapped to a macroblock of the picture 404. In this example, the macroblocks 405, 406, 409, and 410 map or correspond to the macroblock 413 while the macroblocks 407, 408, 411, and 412 map or correspond to the macroblock 414.

[044] After the macroblocks of picture 402 have been mapped to the picture 404, the macroblock type for the macroblocks 413 and 414 is ascertained. It is necessary to

WORKMAN, NYDEGGER & SEELEY
A PROFESSIONAL CORPORATION
ATTORNEYS AT LAW
1000 EAGLE GATE TOWER
60 EAST SOUTH TEMPLE
SALT LAKE CITY, UTAH 84111

determine whether the macroblocks 413 and 414 are intra macroblocks or non-intra macroblocks. This can be determined using the parameters of the original video stream. More specifically, the macroblock types of the original picture are used to determine the macroblock types of the new or transcoded video stream.

[045] In other words, whether the macroblock 413 is an intra or a non-intra macroblock can be determined by applying a weighted mean rounded measure to the macroblocks 405, 406, 409, and 410. For example, each macroblock has a flag that identifies the macroblock type as either intra or non-intra. A 1 represents an intra macroblock type and a 0 represents a non-intra macroblock type. The weighted mean rounded measure is determined as follows. If the macroblocks 405, 406, and 409 have a 1 for their macroblock type flag while the macroblock 410 has a 0 for that flag, then the weighted mean rounded measure for these macroblocks is 1 (round ((1+1+1+0)/4)). Thus, the macroblock 413 in the picture 404 is an intra type macroblock. For those situations where a partial macroblock of the picture 402 is being mapped, then the value of the flag used for these purposes will be weighted accordingly and be between 0 and 1. For example, if half of a macroblock is mapped to a particular macroblock then that macroblock contributes a value of 0.5 to the weighted mean rounded measure.

[046] Determining the macroblock type also requires that other flags including, but not limited to, quant flag, forward flag, backward flag, and pattern flag, be determined. The quant flag indicates whether the quantizer scale of the current macroblock is different from the value currently being used in the decoder and is determined in a similar manner. The pattern flag is also determined in a similar manner. The forward and backward flags indicate whether or not forward and/or backward motion vectors are present in the

macroblock of the picture 404. These flags will be discussed more fully in the next section on motion vector selection.

II. Selection of Motion Vectors and Associated Flags

[047] Motion vectors, in one example, are stored in a 2x2x2 array. The first dimension of the array is associated with top and bottom fields of a picture, the second dimension of the array relates to forward and backward motion, and the third dimension of the array relates to X and Y vectors. Motion vectors are used to obtain data from a reference frame and more specifically identify the location in the reference frame where the data is located.

[048] Figure 4 may also be used to illustrate how to determine motion vectors and associated flags. The motion vectors for the macroblock 413 can be determined using a weighted mean scaled value. Alternatively, the motion type is set to frame motion and the new motion vectors are determined according to how the macroblocks in the picture 402 contributed to the macroblocks in the picture 404. The resulting X and Y values are scaled according to how the picture 404 is scaled or shrunk. After the motion vectors have been determined, a clipping function is employed to ensure that the motion vectors are within appropriate limits.

[049] In another example, the scaled motion vectors of the original macroblocks are used as candidate vectors along with the weighted mean scaled vector. Each of these vectors is evaluated to determine which vector provides a best fit to the data of the resulting video stream. The best fit to the data can be determined, for example, using a goodness of fit measure.

[050] One example of a goodness of fit measure or metric is a subsampled sum of absolute differences (SAD) as a metric as illustrated in Figures 5A and 5B. Figure 5A refers to an

WORKMAN, NYDEGGER & SEELEY
A PROFESSIONAL CORPORATION
ATTORNEYS AT LAW
1000 EAGLE GATE TOWER
60 EAST SOUTH TEMPLE
SALT LAKE CITY, UTAH 84111

original raster scan ordering of pixels from a block. The sequence of numbers in Figure 5B corresponds to a dyadically subdivided selection of the zig zag scan order. Thus, the 0^{th} pixel selected was the 28^{th} in raster scan order and the 31^{st} pixel selected was the 0^{th} in raster scan order. This scan order attempts to get a representation of the entire block. Thus, for a SAD measure, only the pixels in the first row of Figure 5B are used. Alternatively, it is possible to do a subsampled SAD in raster scan order as shown in Figure 5A. The SAD is an example of generating a metric or score that indicates how well a prediction matches the data that is being predicted.

[051] The selection of the motion vectors may also be related to field motion. In one example, the motion can be either frame or field motion, depending on which provides a better fit to the data. The motion vectors are then determined using a weighted mean scaled approach as described above and various settings are evaluated to determine the best fit to the data. Exemplary settings include using field motion plus one of four different settings of the motion vertical field select. Another setting is the frame motion setting.

[052] Fine grain motion estimation may also be performed by evaluating motion vectors in a small search range around, for example, the motion vectors discussed above (weighted means scaled motion vectors, original scaled motion vectors, and field vectors). As the size of the search range is increased, the bit rate of the video stream is typically reduced. However, there may be an increase in the computational and memory requirements related to the increased search range. In another example, it is possible to search for motion vectors independent of previous motion vectors and then compare the newly found motion vectors with the motion vectors obtained as discussed above.

[053] The DCT type flag is a binary value and can be determined per macroblock using the weighted mean rounded procedure previously described for the intra flag. The quantizer

WORKMAN, NYDEGGER & SEELEY
A PROFESSIONAL CORPORATION
ATTORNEYS AT LAW
1000 EAGLE GATE TOWER
60 EAST SOUTH TEMPLE
SALT LAKE CITY, UTAH 84111

scale can be determined using various procedures. The quantizer scale, however, is not limited to values of 0 and 1 like a binary flag. The quantizer scale, therefore, can be determined using various procedures including, but not limited to, the weighted mean rounded procedure, a weighted max rounded procedure, weighted min rounded procedure, a weighted median rounded procedure, and the like. The quantizer scale may be further adjusted according to any rate control mechanism used for the video stream.

[054] The following equations are examples for computing the weighted mean rounded procedure, the weighted max rounded procedure, the weighted min rounded procedure, and the weighted median rounded procedure, given inputs x_i and non-negative weights w_i.

[055] The weighted mean rounded procedure is computed as: Weighted mean rounded = round $((\Sigma_i\ w_i\ x_i)/\ (\Sigma_i\ w_i))$. The weighted min rounded procedure is computed, where i_{min} = arg $\min_i$ (w_i/x_i), as: Weighted min rounded = round(x_{imin}). The weighted max rounded procedure is computed, where i_{max} = arg $\max_i$ $(w_i\ x_i)$, as: Weighted max rounded = round (x_{imax}).

[056] The weighted median rounded procedure is computed as follows: In this procedure, where

h is the highest common factor of the weights w_i (thus $w_i = n_i h$, where n_i is an integer),

(v_k) = a collection of x_i, with each x_i written n_i times,

k_{median} = arg median_k (v_k), and

i_{median} = index corresponding to k_{median},

weighted median round = round ($x_{imedian}$).

[057] The coded block pattern is dependent on the quantization of the DCT coefficients and is computed in a routine manner. The DCT coefficients are computed as illustrated in

WORKMAN, NYDEGGER & SEELEY
A PROFESSIONAL CORPORATION
ATTORNEYS AT LAW
1000 EAGLE GATE TOWER
60 EAST SOUTH TEMPLE
SALT LAKE CITY, UTAH 84111

Figure 6. When a block 602 is a non-intra block, the motion vectors are used to determine the prediction 614 from the reference frames. The prediction 614 is subtracted from the block 602 and a forward DCT 604 is performed on the output. The output is quantized (606) and variable length coded (608) and written as an output bit stream. In the case of I and P frames, inverse quantization 610 and inverse DCT are performed on the output of the quantization process (606) and the prediction is added (614) to new frames as described. Intra blocks do not require prediction and can be quantized without reference to other blocks.

[058] The flags associated with a macroblock, as previously indicated, include a macroblock type flag, a DCT type flag, and a motion type flag. The macroblock flag has 5 bits. Two of these bits are derived from the current state of the transcoder and the remaining flags are determined, in this example, as follows.

[059] Let a_i represents the contribution of macroblocks from the original video stream to the current macroblock of the new video stream. Let $Intra_i$ be a variable that has a value of 1 for intra blocks and a value of 0 for non-intra blocks. Let $Forward_i$ be a variable that has a value of 1 for macroblocks with the "Forward" flag set and a value of 0 otherwise. Let $Backward_i$ be a variable that has a value of 1 for macroblocks with the "Backward" flag set and a value of 0 otherwise. If $((\Sigma_i\ a_i) >= 0.5)$, then the current macroblock is an intra macroblock.

[060] Otherwise, the following steps are taken. Let the Forward flag have a value of 1 if $((\Sigma_i\ a_i\ Forward_i) >= 0.5)$, and have a value of 0 otherwise. If the current macroblock is a B frame, then let the Backward flag have a value of 1 if $((\Sigma_i\ a_i\ Backward_i) >= 0.5)$, and have a value of 0 otherwise. If the current macroblock is a P frame, then the Backward flag is given a value of 0. This example illustrates how to determine the Intra, Forward, and

WORKMAN, NYDEGGER & SEELEY
A PROFESSIONAL CORPORATION
ATTORNEYS AT LAW
1000 EAGLE GATE TOWER
60 EAST SOUTH TEMPLE
SALT LAKE CITY, UTAH 84111

Backward flags for the macroblock type. The remaining flags (quant and pattern) are derived from the quantization scale and the DCT coefficients.

[061] The DCT type for the new macroblock of the new video stream is determined from the DCT $type_i$ of the contributing macroblocks of the original video stream as follows. If $((\Sigma_i\, a_i\, \text{DCT type}_i) >= 0.5)$, then the DCT type of the current macroblock is given a value of 1 and a value of 0 otherwise.

[062] The motion type may be similarly determined where if $((\Sigma_i\, a_i\, \text{Motion type}_i) >= 0.5)$, then the motion type is assigned a value of 1 and a value of 0 otherwise. If fine grained motion estimation is used to determine the motion type, then the motion type and the motion vector are selected based on whether the lowest score comes from the field motion or the frame motion.

[063] The present invention extends to both methods and systems for transcoding a video stream. The embodiments of the present invention may comprise a special purpose or general-purpose computer including various computer hardware, as discussed in greater detail below.

[064] Embodiments within the scope of the present invention also include computer-readable media for carrying or having computer-executable instructions or data structures stored thereon. Such computer-readable media can be any available media that can be accessed by a general purpose or special purpose computer. By way of example, and not limitation, such computer-readable media can comprise RAM, ROM, EEPROM, CD-ROM or other optical disk storage, magnetic disk storage or other magnetic storage devices, or any other medium which can be used to carry or store desired program code means in the form of computer-executable instructions or data structures and which can be accessed by a general purpose or special purpose computer. When information is transferred or provided

WORKMAN, NYDEGGER & SEELEY
A PROFESSIONAL CORPORATION
ATTORNEYS AT LAW
1000 EAGLE GATE TOWER
60 EAST SOUTH TEMPLE
SALT LAKE CITY, UTAH 84111

over a network or another communications connection (either hardwired, wireless, or a combination of hardwired or wireless) to a computer, the computer properly views the connection as a computer-readable medium. Thus, any such connection is properly termed a computer-readable medium. Combinations of the above should also be included within the scope of computer-readable media. Computer-executable instructions comprise, for example, instructions and data which cause a general purpose computer, special purpose computer, or special purpose processing device to perform a certain function or group of functions.

[065] Figure 7 and the following discussion are intended to provide a brief, general description of a suitable computing environment in which the invention may be implemented. Although not required, the invention will be described in the general context of computer-executable instructions, such as program modules, being executed by computers in network environments. Generally, program modules include routines, programs, objects, components, data structures, etc. that perform particular tasks or implement particular abstract data types. Computer-executable instructions, associated data structures, and program modules represent examples of the program code means for executing steps of the methods disclosed herein. The particular sequence of such executable instructions or associated data structures represents examples of corresponding acts for implementing the functions described in such steps.

[066] Those skilled in the art will appreciate that the invention may be practiced in network computing environments with many types of computer system configurations, including personal computers, hand-held devices, multi-processor systems, microprocessor-based or programmable consumer electronics, network PCs, minicomputers, mainframe computers, and the like. The invention may also be practiced in distributed computing

WORKMAN, NYDEGGER & SEELEY
A PROFESSIONAL CORPORATION
ATTORNEYS AT LAW
1000 EAGLE GATE TOWER
60 EAST SOUTH TEMPLE
SALT LAKE CITY, UTAH 84111

environments where tasks are performed by local and remote processing devices that are linked (either by hardwired links, wireless links, or by a combination of hardwired or wireless links) through a communications network. In a distributed computing environment, program modules may be located in both local and remote memory storage devices.

[067] With reference to Figure 1, an exemplary system for implementing the invention includes a general purpose computing device in the form of a conventional computer 20, including a processing unit 21, a system memory 22, and a system bus 23 that couples various system components including the system memory 22 to the processing unit 21. The system bus 23 may be any of several types of bus structures including a memory bus or memory controller, a peripheral bus, and a local bus using any of a variety of bus architectures. The system memory includes read only memory (ROM) 24 and random access memory (RAM) 25. A basic input/output system (BIOS) 26, containing the basic routines that help transfer information between elements within the computer 20, such as during start-up, may be stored in ROM 24.

[068] The computer 20 may also include a magnetic hard disk drive 27 for reading from and writing to a magnetic hard disk 39, a magnetic disk drive 28 for reading from or writing to a removable magnetic disk 29, and an optical disk drive 30 for reading from or writing to removable optical disk 31 such as a CD-ROM or other optical media. The magnetic hard disk drive 27, magnetic disk drive 28, and optical disk drive 30 are connected to the system bus 23 by a hard disk drive interface 32, a magnetic disk drive-interface 33, and an optical drive interface 34, respectively. The drives and their associated computer-readable media provide nonvolatile storage of computer-executable instructions, data structures, program modules and other data for the computer 20. Although the exemplary environment described herein employs a magnetic hard disk 39, a removable magnetic disk 29 and a

removable optical disk 31, other types of computer readable media for storing data can be used, including magnetic cassettes, flash memory cards, digital versatile disks, Bernoulli cartridges, RAMs, ROMs, and the like.

[069] Program code means comprising one or more program modules may be stored on the hard disk 39, magnetic disk 29, optical disk 31, ROM 24 or RAM 25, including an operating system 35, one or more application programs 36, other program modules 37, and program data 38. A user may enter commands and information into the computer 20 through keyboard 40, pointing device 42, or other input devices (not shown), such as a microphone, joy stick, game pad, satellite dish, scanner, or the like. These and other input devices are often connected to the processing unit 21 through a serial port interface 46 coupled to system bus 23. Alternatively, the input devices may be connected by other interfaces, such as a parallel port, a game port or a universal serial bus (USB). A monitor 47 or another display device is also connected to system bus 23 via an interface, such as video adapter 48. In addition to the monitor, personal computers typically include other peripheral output devices (not shown), such as speakers and printers.

[070] The computer 20 may operate in a networked environment using logical connections to one or more remote computers, such as remote computers 49a and 49b. Remote computers 49a and 49b may each be another personal computer, a server, a router, a network PC, a peer device or other common network node, and typically include many or all of the elements described above relative to the computer 20, although only memory storage devices 50a and 50b and their associated application programs 36a and 36b have been illustrated in Figure 1. The logical connections depicted in Figure 1 include a local area network (LAN) 51 and a wide area network (WAN) 52 that are presented here by way of

WORKMAN, NYDEGGER & SEELEY
A PROFESSIONAL CORPORATION
ATTORNEYS AT LAW
1000 EAGLE GATE TOWER
60 EAST SOUTH TEMPLE
SALT LAKE CITY, UTAH 84111

example and not limitation. Such networking environments are commonplace in office-wide or enterprise-wide computer networks, intranets and the Internet.

[071] When used in a LAN networking environment, the computer 20 is connected to the local network 51 through a network interface or adapter 53. When used in a WAN networking environment, the computer 20 may include a modem 54, a wireless link, or other means for establishing communications over the wide area network 52, such as the Internet. The modem 54, which may be internal or external, is connected to the system bus 23 via the serial port interface 46. In a networked environment, program modules depicted relative to the computer 20, or portions thereof, may be stored in the remote memory storage device. It will be appreciated that the network connections shown are exemplary and other means of establishing communications over wide area network 52 may be used.

[072] The present invention may be embodied in other specific forms without departing from its spirit or essential characteristics. The described embodiments are to be considered in all respects only as illustrative and not restrictive. The scope of the invention is, therefore, indicated by the appended claims rather than by the foregoing description. All changes which come within the meaning and range of equivalency of the claims are to be embraced within their scope.

What is claimed and desired to be secured by United States Letters Patent is:

WORKMAN, NYDEGGER & SEELEY
A PROFESSIONAL CORPORATION
ATTORNEYS AT LAW
1000 EAGLE GATE TOWER
60 EAST SOUTH TEMPLE
SALT LAKE CITY, UTAH 84111